

# Effects of time dependency and efficiency on information flow in financial markets

Cheoljun Eom,<sup>1</sup> Woo-Sung Jung,<sup>2</sup> Sunghoon Choi,<sup>1</sup> Gabjin Oh,<sup>3</sup> and Seunghwan Kim<sup>3,4</sup>

*<sup>1</sup>Division of Business Administration,*

*Pusan National University, Busan 609-735, Republic of Korea*

*<sup>2</sup>Center for Polymer Studies and Department of Physics,*

*Boston University, Boston, MA 02215, USA*

*<sup>3</sup>Department of Physics, Pohang University of Science and Technology,*

*Pohang 790-784, Republic of Korea*

*<sup>4</sup>Asia Pacific Center for Theoretical Physics,*

*Pohang 790-784, Republic of Korea*

(Dated: December 2, 2008)

## Abstract

We investigated financial market data to determine which factors affect information flow between stocks. Two factors, the time dependency and the degree of efficiency, were considered in the analysis of Korean, the Japanese, the Taiwanese, the Canadian, and US market data. We found that the frequency of the significant information decreases as the time interval increases. However, no significant information flow was observed in the time series from which the temporal time correlation was removed. These results indicated that the information flow between stocks evidences time-dependency properties. Furthermore, we discovered that the difference in the degree of efficiency performs a crucial function in determining the direction of the significant information flow.

PACS numbers: 89.65.Gh, 05.45.Tp, 89.70.Cf

## I. INTRODUCTION

Recently, researchers have become interested in the information flow occurring between financial assets or markets in an effort to understand the nature of the interaction between assets and the pricing mechanism in markets [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22]. The relationship between spots and derivatives has represented the normal course of study, particularly the manner in which the derivatives that transact with future prices affect the spots [1, 2]. The information flow from developed markets to emerging markets is also an issue in which active research is being conducted [3, 4, 5, 6, 7]. In addition, the information flow with regard to synchronization, integration and segmentation between financial markets by internal and external events has been assessed [8, 9]. Previous studies have attempted to analyze financial data using statistical method including the Granger causality test, the VAR (vector-autoregressive) model, and the GARCH (generalized autoregressive conditional heteroskedasticity) type [10, 11, 12]. However, studies regarding the factors that significantly affect information flow have proven insufficient. Therefore, we have attempted to determine empirically which factors are crucial to the information flow, considering particularly the following factors: the time-dependency property, and differences in the degree of efficiency.

According to the results of previous studies, the financial time series is time-dependent, and the time sequence exerts a significant effect on the information flow. That is to say, in financial markets, there exist many internal and external events which, as time passes, induce price changes via the interactions between stocks at the times that these events occur. In other words, the time scale of return performs a crucial function in the information flow. We have noted that the time scale of return corresponds to the time intervals, particularly when the prices are converted into the returns. Also, the efficiency of information is crucial to the pricing mechanism [23]. The price change in a given individual stock differs from that of others, even though the same information is both instantaneously and fully reflected. This suggests that the degree of efficiency differs for each stock. Therefore, the degree of efficiency significantly affects the information flow. The efficiency we assessed in this study is based on the weak-form efficient market hypothesis (EMH), which assumes that the similarity of past price change patterns are useful in predicting future price changes. We have utilized the approximate entropy (ApEn) method in order to observe the randomness

in the time series [24]. The ApEn method quantitatively calculates complexity, randomness, and prediction power. As the frequency of similarity patterns in the price changes is high, both the randomness and the ApEn remain low. Previous studies have argued that the ApEn evidences significant information by which the degree of efficiency can be measured [25, 26, 27], and the predictive power and ApEn correlated negatively [28].

We have investigated individual stocks traded in the stock markets of Korea, Japan, Taiwan, Canada, and the USA. The entire interactions between stocks are considered, such that the number of interactions is  $N(N - 1)/2$ , where  $N$  is the number of stocks. This technique provides sufficient information flow in the market, allowing us to discover the characteristics of information flow within the context of the whole market. We detected a negative relationship between the time scale of return and the frequency of significant information flow, which supports the notion that the information flow between stocks evidences a time-dependency property. Also, we discovered that the difference in the degree of efficiency between stocks performs a crucial function in determining the direction of the information flow.

In the next section, we describe the data and the methods of the test procedures employed herein. In section III, we present the results obtained in accordance with our established research aims. Finally, we have summarized the findings and conclusions of this study.

## II. DATA AND METHODS

### A. Data

We have assessed the daily closing prices of individual stocks traded in Korea, Japan, Taiwan, Canada, and the USA over a period of 15 years, from January 1992 to December 2006. However, we have excluded the industries which include four individual stocks or less, in order to obtain sufficient statistical analytical features. Therefore, we utilized the daily prices of 95 stocks listed on the KOSPI 200 market index of the Korean stock market, 175 stocks traded in the Nikkei 225 of the Japanese stock market, 132 stocks on the Taiwanese stock market index, 67 stocks on the TSX of the Canadian stock market, and 359 stocks in the S&P 500 market index of the American stock market. In addition, in order to observe sufficient information flow between stocks, we considered the whole links between stocks, in accordance with the formula . The numbers of whole links between stocks for each country

are as follows: 4,465 ( $N = 95$ ) links for Korea, 15,225 ( $N = 175$ ) links for Japan, 8,646 ( $N = 132$ ) links for Taiwan, 2,211 ( $N = 75$ ) links for Canada and 64,261 ( $N = 359$ ) links for the USA. The returns,  $R(t) = \ln(P(t)) - \ln(P(t-1))$ , are calculated using the logarithmic change of the price,  $P(t)$ , in which is the stock price at  $t$  day.

## B. The Granger causality model in the information flow

We investigated the empirical evidence in an effort to determine which factors affect the information flow between stocks. In order to find these factors, we have established the three following steps. First, we calculated the various returns according to the changes in the time scale; second, we calculated returns in which the various time scales apply to the causality model, considering the changes of lag length of the past data as independent variables of the model; and third, we evaluated the observed results of the information flow.

In the first step, we created a return series in accordance with the changes in the time scale, as considering the time-dependency property as a possible factor. The time scale  $k$  refers to the time intervals, when the prices are converted into the returns. The time interval varies from 1 day to 5 days ( $k = 1, \dots, 5$ ). The return  $R_k$  corresponds to the time scale  $k$  and is calculated by  $R_k = \sum_{t=1}^k R(t)$ .

In the second step, we employed the Granger causality model to determine the direction of the significant information flow. In this model, we determined the lag length  $l$  of the past data as independent variables,  $X_{t-l}$  and  $Y_{t-l}$ , in order to explain the current price changes as the dependent variables,  $X_t$  and  $Y_t$ . Using each return with various time scales, we have assessed the significant information flows as  $X \Rightarrow Y$  and/or  $Y \Rightarrow X$ , considering the lag length of the past data change,  $l = 1, \dots, 5$ , which can be defined as:

$$X \Rightarrow Y : \quad Y_t = C + \sum_{l=1}^L \alpha_l Y_{t-l} + \sum_{l=1}^L \beta_l X_{t-l} \quad (1)$$

(The null hypothesis  $H_0$  : X does not Granger-cause Y)

and

$$Y \Rightarrow X : \quad X_t = C^* + \sum_{l=1}^L \alpha_l^* X_{t-l} + \sum_{l=1}^L \beta_l^* Y_{t-l} \quad (2)$$

(The null hypothesis  $H_0$  : Y does not Granger-cause X)

where  $C$ ,  $\alpha$  and  $\beta$  were the estimated coefficients, the minimum and maximum of the lag length. We utilized Granger F-statistics (the Wald statistics) for each lag length with a statistical significance level of 5% in order to determine the degree of significant information flow. Considering both the time scale and the lag length, we have statistically classified the significant information flow into three types. If both Eq. 1 and Eq. 2 are significant, the information flow results in the mutual exchange of information,  $X \Leftrightarrow Y$ . When either Eq. 1 or Eq. 2 is significant, we observe a one-way direction of information,  $X \Rightarrow Y$  or  $Y \Rightarrow X$ . As both Eq. 1 and Eq. 2 are insignificant, there is no exchange of information, which can be expressed as  $X \parallel Y$ .

In the third step, we created measurements to evaluate the significant information flow observed in the second step since we considered both the time scale ( $k = 1, \dots, 5$ ) and the lag length ( $l = 1, \dots, 5$ ) in the model. Therefore, we utilized the frequency ratio (FR) of the significant information flow, which is defined as:

$$\text{FR} = \frac{F_q}{N(N-1)/2} \quad (3)$$

where  $F_q$  represents the frequency of significant information flow for each of the three cases;  $X \Leftrightarrow Y$ ,  $X \Rightarrow Y$  or  $Y \Rightarrow X$ , and  $X \parallel Y$ . As we investigated the changes in the information flow as functions of the time scale, we were able to confirm whether or not the information flow evidenced the time-dependency properties.

### C. The approximate entropy

The ApEn was utilized to observe the randomness in the time series so that we could quantitatively calculate complexity, randomness, and prediction power. The ApEn also assesses the degree of efficiency. The ApEn is defined as

$$\text{ApEn}(m, r) = \Phi^m(r) - \Phi^{m+1}(r) \quad (4)$$

where  $m$  represents the embedding dimension and  $r$  is the tolerance of similarity. The  $\Phi^m(r)$  is given by

$$\Phi^m(r) = \frac{1}{(N-m+1)} \sum_{t=1}^{(N-m+1)} \ln [C_i^m(r)] \quad (5)$$

where  $C_i^m = B_i(r)/(N - m + 1)$  and  $B_i(r)$  is the number of data pairs within the tolerance of similarity  $r$ . Additionally, we calculated the similarity in the time series of each price change pattern  $(u(k), k = 1, \dots, m)$  by the distance  $d[x(i), x(j)]$  between two vectors,  $x(i)$  and  $x(j)$  defined by

$$B_i = d[x(i), x(j)] \leq r \quad (6)$$

and

$$d[x(i), x(j)] = \max_{k=1, \dots, m} (|u(i + k - 1) - u(j + k - 1)|). \quad (7)$$

In the above equations, the ApEn compares the relative magnitude between repeated pattern occurrences for the embedding dimensions  $m$  and  $m + 1$ . The ApEn becomes smaller as the frequency of similar price change patterns for the embedding dimension  $m$  becomes equal to those for the embedding dimension  $m + 1$ . When both frequencies are equal, the ApEn is 0. Therefore, the ApEn is small when the frequency of the similar price change pattern is high. The time series data evidences a low degree of randomness, and the efficiency of the financial time series becomes low. In this study, we have utilized the embedding dimension  $m = 2$  and the tolerance of similarity  $r = 20\%$  of the standard deviation of the time series.

### III. RESULTS

#### A. The time-dependency property

We attempted to determine whether the information flow changes with the time scale. Fig. 1 shows the FR for 175 stocks listed on the Nikkei 225 of the Japanese stock market. The lag length  $l$  and the time scale  $k$  were shown to vary from 1 day to 5 days. The shuffled data was analyzed in order to eliminate the temporal correlation in the time series. The shuffled data evidences the same statistical properties, including the mean, variance, skewness, and kurtosis with the original time series. In Figs. 1(a) and (b), the X-axis denotes the time scale, whereas the Y-axis refers to the of the information flow. The X-axes of Figs. 1(c) and (d) correspond to the lag length. The circles, squares, and triangles indicate the mutual exchanges, the one-way direction, and no exchanges of the information, respectively. Figs. 1(a) and (c) represent the results from the original data, and Figs. 1(b) and (d) show those for the shuffled data.

In Fig. 1, we determined that the information flow observed by the Granger causality model is influenced significantly by the time scale. The FRs for the mutual exchange of information and the one-way direction of information decrease as the time scale  $k$  increases from 1 day to 5 days, whereas that for no exchange of information increases. However, we have not discovered any significant trend as the lag length  $l$  increases. This means that the information flow evidences a time-dependency property. Fig. 2 represents the as a function of for other markets, including Korea, Taiwan, Canada, and the USA. Figs. 2(a) and (b) show the one-way direction of information, and (c) and (d) represent the mutual exchange of information.

Figs. 1(b), (d) and 2(b), (d) correspond to the FRs for the shuffled data with no time correlations. In those figures, we determined that the information flow is quite sensitive to the time sequence. The FRs for the mutual exchange and the one-way direction of information is very small, regardless of the values of  $k$  and  $l$ . However, we noted that the FR for no exchange of information was higher in Figs. 1(b) and (d). Therefore, we were unable to detect any significant information flow in cases in which the time sequence was removed.

## B. The difference in the degree of efficiency

Next, we attempted to determine empirically whether information flow between stocks was affected by differences in the degree of efficiency between stocks. The price reaction of individual stocks differed in accordance with whether the new information was instantaneously and fully reflected in the price. The fact is, the degree of efficiency differs among each stocks. Therefore, the difference in the degree of efficiency between stocks may influence the interactions between stocks, and may exert a significant effect on information flow. We have tested the assumption that the more efficient stocks reacts quickly against the information, such that the price also changes more quickly than does the less efficient stock. Also, the more efficient stock with relatively more sufficient information delivers the information to the less efficient stock. Finally, the information flow is correlated with the difference in the degree of efficiency.

In order to quantify the degree of efficiency, we utilized the ApEn method, which quantifies the degree of complexity, irregularity, and randomness in the time series. The ApEn is

calculated on the basis of the degree of similarity patterns in the time series. Therefore, the ApEn is related to the EMH, and can thus be used to quantify the degree of efficiency in the information flow. If the time series evidences a high degree of randomness and complexity, the ApEn is large, whereas the ApEn of the random time series is quite high.

Fig. 3 demonstrates the frequency ratios for  $X^E \Rightarrow Y^E$  (triangle up) and  $Y^E \Rightarrow X^E$  (triangle down) efficient stocks when  $X^E$  is the more efficient stock as compared with  $Y^E$ . The one-way direction of information cases (circle) from Figs. 1 and 2 are provided for purposes of comparison. In Fig. 3, the FR for  $X^E \Rightarrow Y^E$  was shown to be higher than that for the opposite direction  $Y^E \Rightarrow X^E$  for all of the analyzed markets. This means that the more efficient stock  $X^E$  exerts the driving force of the information flow into the less efficient stock  $Y^E$ . In particular, the difference between  $X^E \Rightarrow Y^E$  and  $Y^E \Rightarrow X^E$  for the Korean and Canadian markets are relatively obvious and clear. However, the reason why those markets' features are more clear remains an open question, and will be the subject of our future work. On the basis of these results, we discovered that the difference in the degree of relative efficiency between stocks influences the direction of the information flow, and thus plays an important factor.

#### IV. CONCLUSIONS

We have investigated stock market data in order to determine which factors perform crucial functions in the information flow between stocks. The time dependency property and the difference in the degree of efficiency were assessed using the Granger causality model and the approximate entropy.

We found that the information flow observed via the Granger causality model was influenced significantly by the time scale. The time scale and the frequency of the significant information flow were correlated negatively. However, we were not able to detect a significant information flow when the temporal time correlation was removed from the time series. We also discovered that the difference in the degree of efficiency played a crucial role in determining the direction of the information flow. We confirmed that the information flowed from the more efficient stock to the less efficient stock. These findings indicated that the information flow evidences a time-dependency property, and is influenced by the difference in the degree of efficiency. In our opinion, when researchers further assess the information



flow between assets or markets, they should consider the time-dependency property and the difference in the degree of efficiency.

---

- [1] L. M. Han, L. Misra, J. Futures markets **10**, 273 (1990).
- [2] N. T. Milonas, J. Futures Markets **6**, 443 (1986).
- [3] O. F. Aydi, U. B. Dufrene, A. Chatterjee, Int. Rev. Finan. Anal. **7**, 83 (1998).
- [4] Y. Hamao, R. Masulis, V. Ng, Rev. Financial Studies **3**, 281 (1990).
- [5] M. King, S. Sadhwani, Rev. Financial Studies **3**, 5 (1990).
- [6] W.-S. Jung, S. Chae, J.-S. Yang, H.-T. Moon, Physica A **361**, 263 (2004).
- [7] W.-S. Jung, O. Kwon, F. Wang, T. Kaizoji, H.-T. Moon, H. E. Stanley, Physica A **387**, 537 (2008).
- [8] D. Panton, V. Lessig, O. Joy, J. Financ. Quant. Anal. **11**, 415 (1976).
- [9] S. Maslov, Physica A **301**, 397 (2001).
- [10] C. W. Granger, Econometrica **37**, 424 (1969).
- [11] R. Engel, C. W. Granger, Econometrica **55**, 252 (1987).
- [12] C. Eun, S. Shim, J. Financ. Quant. Anal. **24**, 241 (1989).
- [13] C. E. Shannon, W. Weaver, *The Mathematical Theory of Information* (University of Illinois Press, 1994).
- [14] R. N. Mantegna, Eur. Phys. J. B **11**, 193 (1999).
- [15] V. Plerou, P. Gopikrishnan, B. Resnow, L.A. Nunes Amaral, H. E. Stanley, Phys. Rev. Lett. **83**, 1471 (1999).
- [16] T. Schreiber, Phys. Rev. Lett. **85**, 416 (2000).
- [17] R. Marschinski, H. Kantz, Eur. Phys. J. B **30**, 275 (2002).
- [18] L. Kullmann, J. Kertész, K. Kaski, Phys. Rev. E **66**, 026125 (2002).
- [19] J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertész, A. Kanto, Phys. Rev. E **68**, 056110 (2003).
- [20] B. Tóth, J. Kertész, Physica A **360**, 505 (2006).
- [21] R. Kumar, S. Sinha, Phys. Rev. E **76**, 046116 (2007).
- [22] O. Kwon, J.-S. Yang, Physica A (2008), doi:10.1016/j.physa.2008.01.007
- [23] E. F. Fama, J. Finance **25**, 383 (1970).
- [24] S. M. Pincus, Proc. Natl. Acad. Sci. USA **88**, 2297 (1991).

- [25] S. M. Pincus, R. E. Kalman, Proc. Natl. Acad. Sci. USA **101**, 13709 (2004).
- [26] T. Kim, C. Eom, G. Oh, Korean J. Finance **18**, 239 (2005).
- [27] G. Oh, S. Kim, C. Eom, Physica A **382**, 209 (2007).
- [28] C. Eom, G. Oh, W.-S. Jung, arxiv.org: 0708.4178.

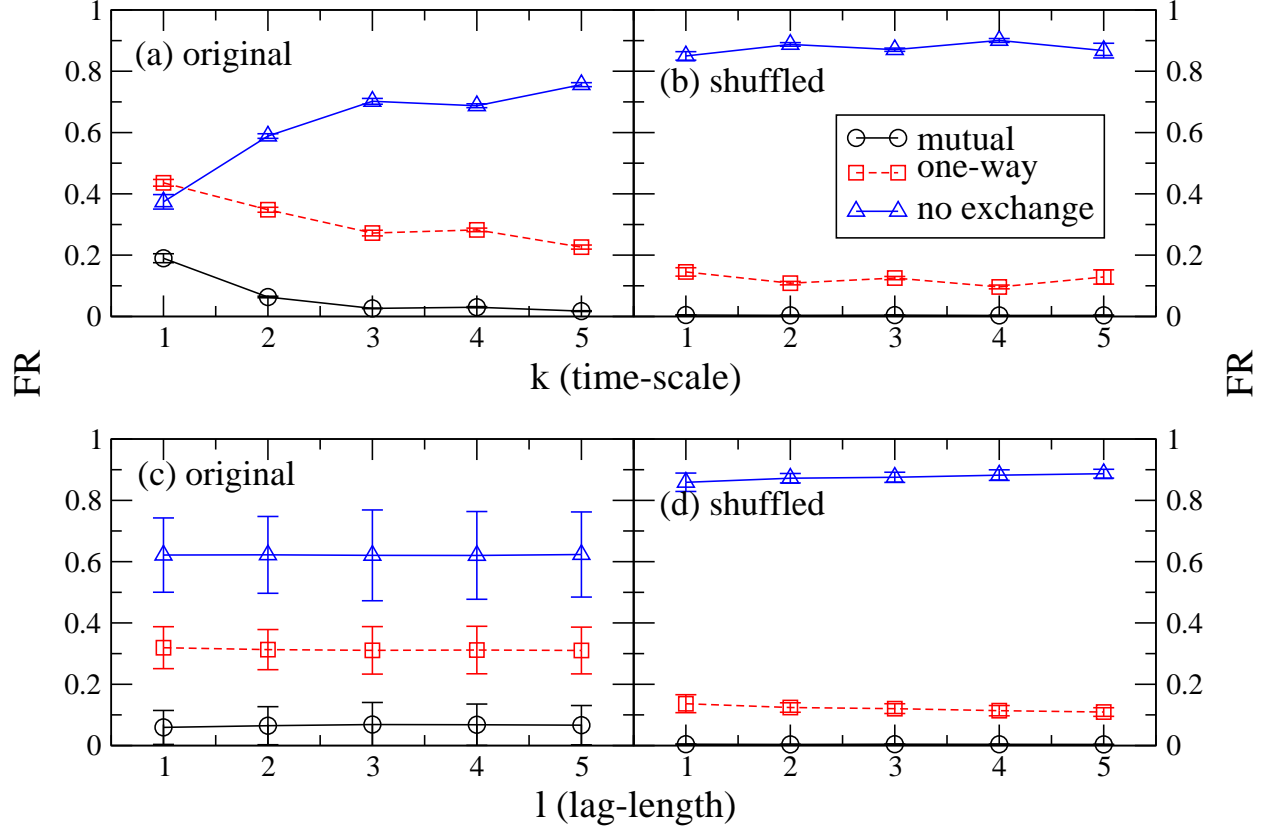


FIG. 1: (Color online). The frequency ratio (FR) as functions of the time scale  $k$  and the lag length  $l$  for the Japanese market. Both  $k$  and  $l$  range between 1 and 5. The original data was analyzed in (a) and (c), and the shuffled data in (b) and (d). The circle corresponds to the mutual exchange, the square to the one-way direction exchange, and the triangle to no exchange, respectively.

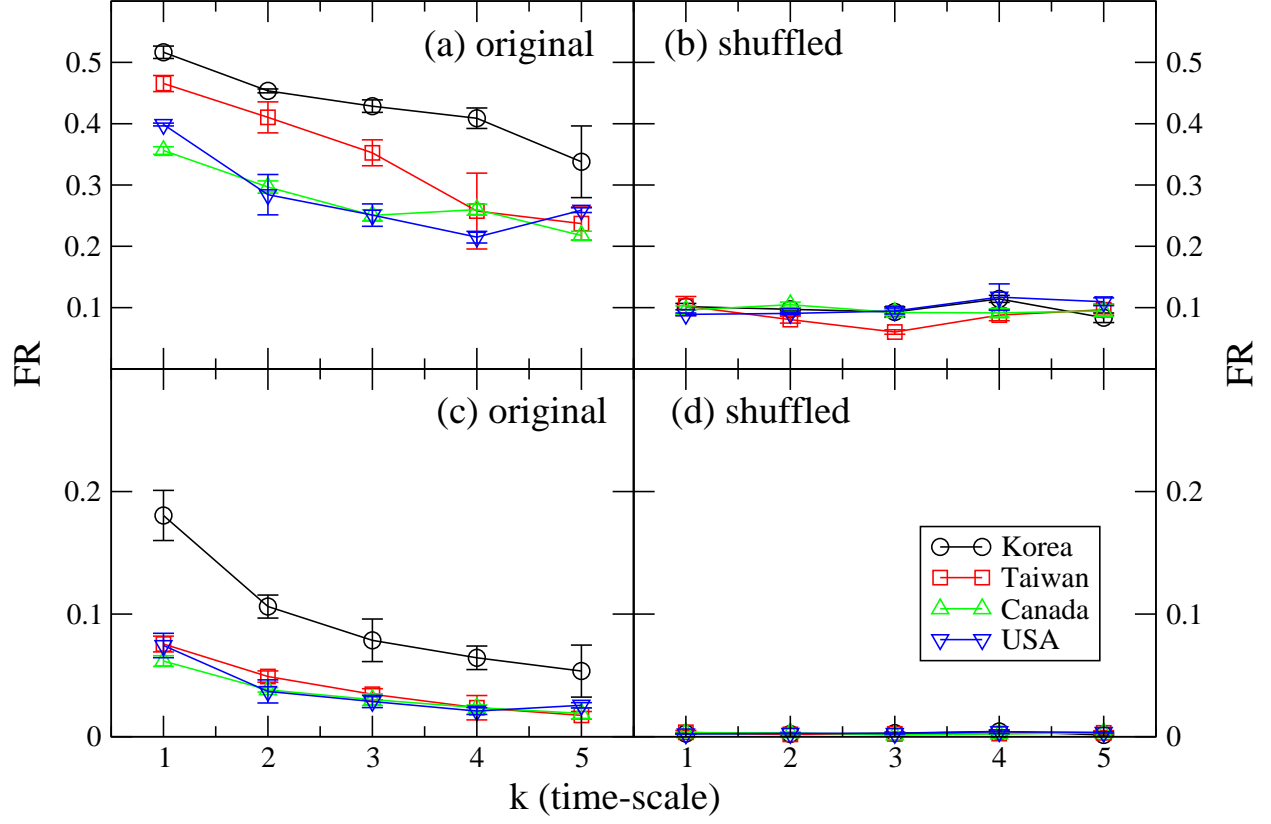


FIG. 2: (Color online). The frequency ratio (FR) as functions of the time scale  $k$  for the Korean (circle), Taiwanese (square), Canadian (triangle up), and US markets (triangle down). The original data was assessed in (a) and (c), and the shuffled data in (b) and (d). (a) and (b) represent the one-way direction exchange, and (c) and (d) the mutual exchange.

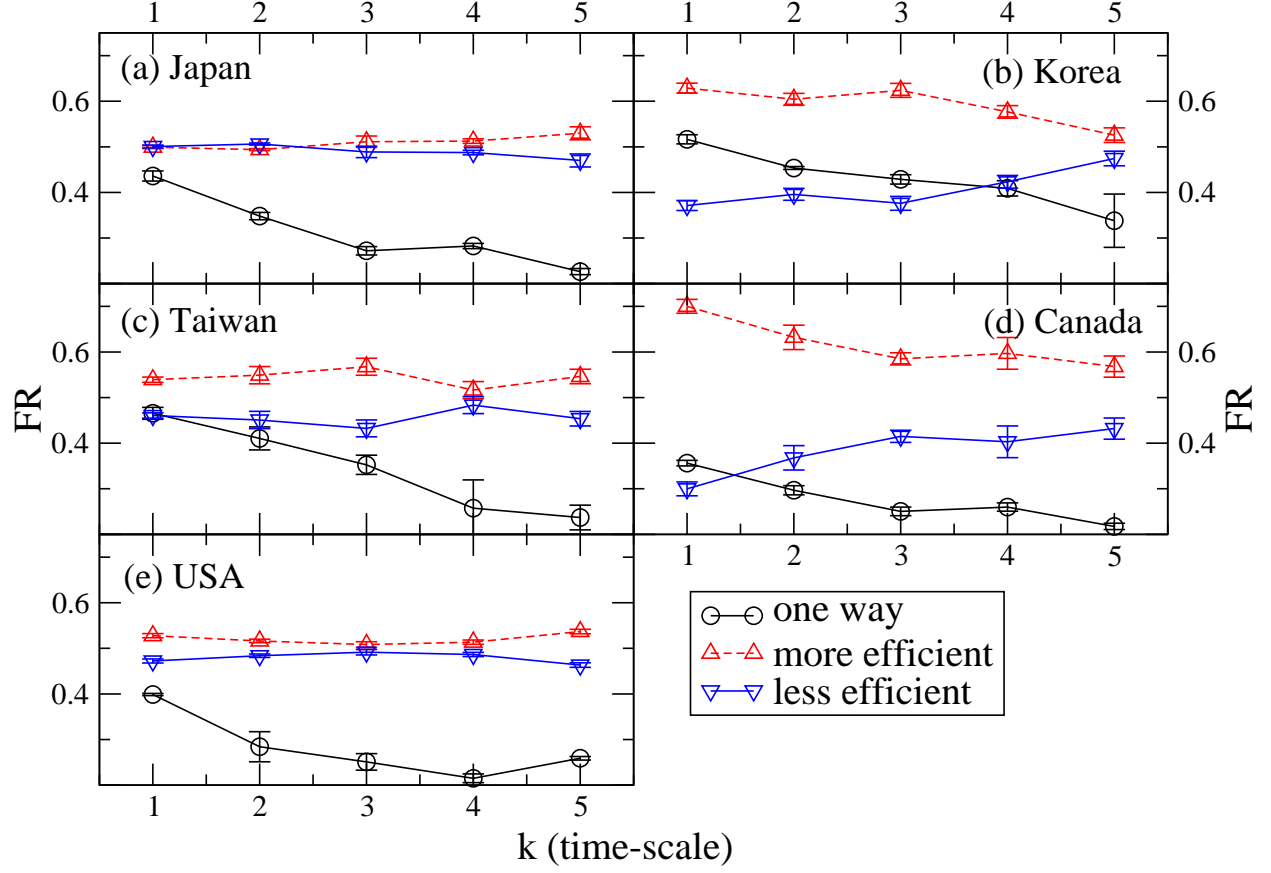


FIG. 3: (Color online). The frequency ratio (FR) for  $X^E \Rightarrow Y^E$  (triangle up) and  $Y^E \Rightarrow X^E$  (triangle down) efficient stocks when  $X^E$  is the more efficient stock in comparison with  $Y^E$ . The one-way direction case (circle) is provided from Figs. 1 and 2 for purposes of comparison.